

SspI

GCAC TTTTCGGG GAAATGTGCGCGGAACCCCTATTTGTTTATTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCGTATAAATGCTTCAATAATATTGAAAAAGGA  
CGTGAAAAGCCCCTTTACACGCGCCTTGGGGATAAAACAATAAAAAGATTTATGTAAGTTTATACATAGGCGAGTACTCTGTTATTGGGACTATTTACGAAGTTATTATAACTTTTTCTCT

H F S G K C A R N P Y L F I F L N T F K Y V S A H E T I T L I N A S I I L K K E

AGAGTATGAGTATTC AACATTTCCGTGTCGCCCTTATTCCTTTTTGCGGCATTTTGCCTTCCTGTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGG  
TCTCATACTCATAAGTTGTAAGGCACAGCGGAATAAGGGAAAAACGCCGTA AAAACGGAAGGACAAAAACGAGTGGGCTTTGCGACCAC TTTTCA TTTTCTACGACTTCTAGTCAACC

E Y E Y S T F P C R P Y S L F C G I L P S C F C S P R N A G E S K R C . R S V G

GTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTGCGCCCGAAGAACGTTTTCCAATGATGAGCAC TTTTAAAGTTCTGCTATGTGGCGCGGTAT  
CACGTGCTACCCCAATGTAGCTTGACCTAGAGTTGTCGCCATTCTAGGAAGCTCTCAAAGCGGGGCTTCTTGCAAAGGTTACTACTCGTGAAAAATTTCAAGACGATACACCGGCCATA

C T S G L H R T G S Q Q R . D P . E F S P R R T F S N D E H F . S S A M W R G I

Scal

TATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCATACACTATTCTCAGAATGACTTGGTTGAGTACTACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAG  
ATAGGGCATAACTGCGGCCCGTTCTCGTTGAGCCAGCGCGTATGTGATAAGAGTCTTACTGAACCAACTCATGAGTGGTCAGTGTCTTTTCGTAGAAATGCCTACCGTACTGTCATTCTC

I P Y . R R A R A T R S P H T L F S E . L G . V L T S H R K A S Y G W H D S K R

PvuI

AATTATGCAGTGCTGCCATAACCATGAGTGATAAACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTGCACAACATGGGGATCATGTAAGTC  
TTAATACGTCACGACGGTATTGGTACTCACTATTGTGACGCCGGTTGAATGAAGACTGTTGCTAGCCTCCTGGCTTCTCGATTGGCGAAAAAACGTTGTGATCCCTAGTACATTGAG

I M Q C C H N H E . . H C G Q L T S D N D R R T E G A N R F F A Q H G G S C N S

GCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAAC AACGTTGCGCAAAC TATTAAC TGCGGAAC TACTTACTC  
CGGAAGTAGCAACCTTGGCCTCGACTTACTTCGGTATGGTTTGTGCTCGCACTGTGGTGTACGGACATCGTTACCGTTGTTGCAACGCGTTTGATAATTGACCGCTTGATGAATGAG

P . S L G T G A E . S H T K R R A . H H D A C S N G N N V A Q T I N W R T T Y S

TAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCAC TCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTG  
ATCGAAGGGCCGTTGTTAATTATCTGACCTACCTCCGCTATTTCAACGCTCTGGTGAAGACGCGAGCCGGGAAGGCCGACCAAAATAACGACTATTTAGACCTCGGCCACTCGCAC

S F P A T I N R L D G G G . S C R T T S A L G P S G W L V Y C . . I W S R . A W

GGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAG  
CCAGAGCGCCATAGTAACGTCGTGACCCCGGTCTACCATTCGGGAGGGCATAGCATCAATAGATGTGCTGCCCTCAGTCCGTTGATACCTACTTGCTTTATCTGTCTAGCGACTCTATC

V S R Y H C S T G A R W . A L P Y R S Y L H D G E S G N Y G . T K . T D R . D R

GTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACCTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATC 1080  
CACGGAGTGACTAATTCGTAACCAATTGACAGTCTGGTTCAAATGAGTATATATGAAATCTAACTAAATTTTGAAGTAAAAATTAATTTTCTAGATCCACTTCTAGGAAAAACTATTAG  
C L T D . A L V T V R P S L L I Y T L D . F K T S F L I . K D L G E D P F . . S  
TCATGACCAAAATCCCTTAACGTGAGTTTTCTGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAA 1200  
AGTACTGGTTTTAGGGAATTGCACCTCAAAGCAAGGTGACTCGCAGTCTGGGGCATCTTTCTAGTTTCTAGAAGAACTCTAGGAAAAAAGACGCGCATTAGACGACGAACGTTTGT  
H D Q N P L T . V F V P L S V R P R R K D Q R I F L R S F F S A R N L L L A N K  
AAAAACCACCGCTACCAGCGGTGTTTTGTTTCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACGGCTTCAGCAGAGCGCAGATACCAAACTGTCCTTCTAGTGTAGCCGT 1320  
TTTTGGTGGCGATGGTCCACCAACAACAGCGCTAGTTCTCGATGGTTGAGAAAAGGCTTCCATTGACCGAAGTCGCTCGCGTCTATGGTTTATGACAGGAAGATCACATCGGCA  
K T T A T S G G L F A G S R A T N S F S E G N W L Q Q S A D T K Y C P S S V A V  
AGTTAGGCCACCCTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGCTTACCGGGTGGACTCAAGAC 1440  
TCAATCCGGTGGTGAAGTTCTTGAGACATCGTGGCGGATGATGGAGCGAGACGATTAGGACAATGGTCACCGACGACGGTCACCGCTATTTCAGCAGCAATGGCCCAACCTGAGTTCTG  
V R P P L Q E L C S T A Y I P R S A N P V T S G C C Q W R . V V S Y R V G L K T  
GATAGTTACCGGATAAGGCGCAGCGGTGGGGCTGAACGGGGGTTCTGTCACACAGCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCG 1560  
CTATCAATGGCTATTCGGCGTCCAGCCGACTTGGCCCCAAGCAGTGTGTCGGTGAACCTCGCTTGCTGGATGTGGCTTGACTCTATGGATGTCGCCTGACTACTCTTTCCG  
I V T G . G A A V G L N G G F V H T A Q L G A N D L H R T E I P T A . A M R K R  
CCACGCTTCCGAAGGGAGAAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGAAACGCCTGGTATCTTTATAGTCTGTCCGGT 1680  
GGTGCGAAGGGTCCCTCTTTCCGCTGTCCATAGGCCATTCCGCTCCAGCCTTGCTCTCGCTGCTCCC TCGAAGTCCCCTTTGCGGACCATAGAATATCAGGACAGCCCA  
H A S R R E K G G Q V S G K R Q G R N R R A H E G A S R G K R L V S L . S C R V  
TTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAAGGGGGCGGAGCCTATGGAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCGCTTTTGTGGCCTTTTGTCTC 1800  
AAGCGGTGGAGACTGAACTCGACGTA AAAACACTACGAGCAGTCCCCCGCTCGGATACCTTTTTGCGGTGCTGCGCGGAAAAATGCCAAGGACCGGAAAACGACCGGAAAACGAG  
S P P L T . A S I F V M L V R G A E P M E K R Q Q R G L F T V P G L L L A F C S  
ACATGTTCTTCTCGCTTATCCCCTGATTCTGTGGATAACCGTATTACCGCTTTGAGTGAGCTGATACCGCTCGCCGACGCCAACGACCGAGCGCAGCGAGTCAGTGAGCGGGAAG 1920  
TGTACAAGAAAGGACGCAATAGGGGACTAAGACACCTATTGGCATAATGGCGAAACCTACTCGACTATGGCGAGCGCGTGGCTTGTGGCTCGCTCGCTCAGTCACTCGCTCCTTC  
H V L S C V I P . F C G . P Y Y R L . V S . Y R S P Q P N D R A Q R V S E R G S

PvuII

CGGAAGAGCGCCCAATACGCAAACCGCTCTCCCGCGCGTTGGCCGATTCAATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGA  
GCCTTCTCGGGGTTATGCGTTTGGCGGAGAGGGGCGCGCAACCGGCTAAGTAATTACGTCGACCGTGTCTGCAAAGGGCTGACCTTTCGCCGTCACCTCGCGTTGCGTTAATTACACT

2040

G R A P N T Q T A S P R A L A D S L M Q L A R Q V S R L E S G Q . A Q R N . C E

GTTAGCTCACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGCCA  
CAATCGAGTGAGTAATCCGTGGGGTCCGAAATGTGAAATACGAAGGCCGAGCATACAACACACCTTAACACTCGCTATTGTTAAAGTGTGTCCTTTGTCGATACTGGTACTAATGCGGT

2160

L A H S L G T P G F T L Y A S G S Y V V W N C E R I T I S H R K Q L . P . L R Q

Acc65I  
KpnI

AGCTCGAAATTAACCCCTCACTAAAGGGAACAAAAGCTGGTACCCCGCATTGATGGTTTCGCCAAAAAGATGTTTTAATAGCAAATTAATACTACACTTTAAAAGGAAAAGGAACGTTC  
TCGAGCTTTAATTGGGAGTGATTTCCCTTGTTCGACCATGGGGCGTAAACTACCAAAGCGGATTTTCTACAAAATTATCGTTAATTATGATGTGAAATTTCCCTTTCCCTTCAAG

2280

A R N . P S L K G T K A G T P H L M V S P K K M F . . Q I N T T L . K E K E R S

EcoRI

PvuII

XbaI

CACTATGCAGTGGAATTCGGTCGATCGACTCTAGAGGATCAGAAAATATCGCCATAAAAGACAGAATAAGTCATCAGCGGTTGTTTCATTTCCTATATTTTTTTTTTTTTTTTTTTT  
GTGATACGTCACCTTAAGCCAGCTAGCTGAGATCTCCTAGTCTTTAATAGCGGTATTTCTGTCTTATTCAGTAGTCGCCAACAAAGTAAAGGATATAAAAAAAAAATAAAAAAAAAA

2400

nmt1(41X) promoter

T M Q W N S V D R L . R I R K L S P . K T E . V I S G C F I S Y I F F L F F Y F

TTTAATAAGGGAAAATTTAACGTCTAAGGATACAGAAGATTGTTAGCACATTAAGTAATAAAGGCTTAAGTAGTAAGTGCCCTAGCATGTTATTGTATTTCAAAGGACATAATCTAAAA  
AAATTATCCCTTTTAAATGTCAGATTCCCTATGTCTTCTAACAATCGTGAATTTCAATTTCCGAATTCATCATTACGGAATCGTACAATAACATAAAGTTTCTGTATTAGATTTT

2520

nmt1(41X) promoter

L I R E N L T S K D T E D C . H I K V I K A . V V S A L A C Y C I S K D I I . N

TAATAACAATATCATTCTCACAAAGTTATTCAATTTCTTTTTTTTTTCTAATAATATCAAGAAATGATTATTTGTTTGACATAAGTCAACTAATTTATTTAATATGCTGGATTAATCTT  
ATTATTGTTATAGTAAGAGTGTTCAATAAGTAAAAAGAAAAAAAGATTATTATAGTCTTACATAATAAACAAACTGATTTCAGTTGATTAAATAAATTATACGACCTAATTAGAA

2640

nmt1(41X) promoter

N N N I I S H K L F N F L F F F . . Y Q E C I I C L T . V N . F I . Y A G L I L

GCAGACATGTA AATTAACAAGTTTTAGTCAAATAACGTTGAAGTTTCAATGAAC TCAATAATTTCTCTTTTTTTTATATAACCATAAGTCTAATCTGATTTATATTTTCCGCAGGGAT 2760  
CGTCTGTACATTTAATTGTTCAA AATCAGTTTATGCAACTTCAAAGTTACTTGAGTTTATTAAGAGAAAAA AATATATTGGTATTTCAGATTAGACTAAATATAAAAAGGCGTCCCTA

nmt1(41X) promoter

Q T C K L T S F S Q I T L K F Q . T Q I I S L F F Y I T I S L I . F I F S A G I

CAACTGAAGTTATGACATTTGGATTGGATCACTTATAACCTTGGTCGCCAAATAATACAAAAATCAGCGTTATAAACAAGAAGGTTTTGTTAAGAAATTAATCCTCTTTCTTGATAA 2880  
GTTGACTTCAACTACTGTAACCTAACCTAGTGAATATTGGAAC CAGCGTTTATTATGTTTTAGTCGCAATATTTGTTTCTCCAAAAACAATTCTTTAATTAGGAGAAAGAACTATT

nmt1(41X) promoter

N . S Y D I W I G S L I T L V A K . Y K N Q R Y K T K K V F V K K L I L F L D K

BstXI

GAAAGTTGAACCGAAATTCAGATACTGATATATGAAAAATACCCACAATTTGGGAATAGCGCAAGCCTCAATTTAAACAATAGGTGAGGACACATGATAATGACCTCAATGATTGT 3000  
CTTTCACCTTGGCTTTAACGTCATGACTATATACTTTTATTATGGGTGTTAAAACCTTATCGCGTTCCGAGTTAAATTTGTTATCCACTCCTGTGTACTATTACTGGAGTTACTAACA

nmt1(41X) promoter

K V E P K L Q I L I Y E N N T H N F G N S A S L N L N N R . G H M I M T S M I V

TAGAAGAAAAGAGCCTCATTACAAAATCGAAAAATGAATGGTTGGGTACAAGTTTCCAAAACATGGTAAAGTGGACTTTGCGTATGAGACGTAATAGAAAAAACACTTGTTATATGTT 3120  
ATCTTCTTTTCTCGGAGTAATGTTTTAGCTTTTACTTACCAACCCATGTTCAAAGGTTTTGTACCATTTACCTGAAACGCATACTCTGCATTTATCTTTTTTTGTGAACAATATACAA

nmt1(41X) promoter

R R K E P H Y K I E K . M V G Y K F P K H G K V D F A Y E T . I E K N T C Y M F

XbaI

TTCTAGAATTATTGTTGTCTCTTTATGGTTGGATGATGCAAAATAGTAATTTCCGGTTAGTTGCTGTAAAACACCACGAGACAAATAGATATGGATATTTATTAATCAGGAAAAACGTAA 3240  
AAGATCTTAATAACAACAGAGAAATACCAACCTACTACGTTTATCATTAAAGCCAATCAACGACATTTGTGGTCTCTGTTATCTATACCTATAAATAATTTAGTCTTTTTTGCATT

nmt1(41X) promoter

S R I I V V S L W L D D A K . . F R L V A V K H H E T N R Y G Y L L N Q E K R N

HindIII

CTCTCGGCTACTGGATGGTTCAGTCACCCAACGATTACTGGGGAGAGAAAAACAGGGCAAAGCTTAAAGGAATCCGATTGTCATTCGGCAATGTGCAGCGAAACTAAAAACCGGA 3360  
GAGAGCCGATGACCTACCAAGTCAGTGGGTTGCTAATGACCCCTCTCTTTGTCCCGTTTTCGTTTTCGAATTTCTTAGGCTAACAGTAAGCCGTTACACGTCGCTTTGATTTTTGGCCT

nmt1(41X) promoter

S R L L D G S V T Q R L L G R E N R A K A K L K G I R L S F G N V Q R N . K P D

TAATGGACCTGTTAATCGAAACATTGAAGATAAAGGAAGAGGAATCCTGGCATATCATCAATTGAATAAGTTGAATTAATTATTTCAATCTCATTCTCACTTTCTGACTTATAGTCGCTT 3480  
ATTACCTGGACAATTAGCTTGTAACTTCTATTTCCTTCTCCTTAGGACCGTATAGTAGTAACTTATTCAACTTAATTAATAAAGTTAGAGTAAGAGTGAAAGACTGAATATCAGCGAA

nmt1(41X) promoter

N G P V N R N I E D K G R G I L A Y H Q L N K L N . L F Q S H S H F L T Y S R F

NdeI Apal

TGTTAAATCATATGGGGCCCTGCTTATCTTTCTTCTCCGAAAGTGCTACTTCAGAAAAGCAAGAAGTGTACGGATAAAGGATAAGTATATCTTCTACTACAATACGACAAGAAC 3600  
ACAATTTAGTATACCCCGGGGACGAATAGAAAAGAAGAAGGCTTTCACGATGAAGTCTTCTCGTTCCTTGACATGCCTATTCCATTCATATAGAAGATAGATGATGTTATGCTGTTCTTG

nmt1(41X) promoter

V K S Y G A P A Y L S S S E S A T S E R A R T V R I R I S I S S I Y Y N T T R T

BstXI

TCTCGAACTCTCCCTCCCCTTACTTTAACTAATAGATTATGCGAGATTTTCGTCAGAGACTTTCGACGGTAAAACCATAAACATTGGAAGTTGAATCTTCCGATACCATCGACAACGTTAA 3720  
AGAGCTTGAGAGGGAGGGTGAATGAAATGATTATCTAATACGTCTAAAAGCAGTTCTGAAACTGGCCATTTGGTATTGTAACCTTCAACTTAGAAGGCTATGGTAGCTGTTGCAATT

ubiquitin

L E L S L P L Y F N . . I M Q I F V K T L T G K T I T L E V E S S D T I D N V K

GTCGAAAATTCAGACAAGGAAGGTATCCCTCCAGATCAACAAGATTGATCTTTGCCGGTAAGCAGCTAGAAGACGGTAGAACGCTGTCTGATTACAACATTCAGAAGGAGTCCACCTT 3840  
CAGCTTTAAGTTCTGTTCCCTCCATAGGGAGGTCTAGTGTGTTCTAACTAGAAAACGCCATTCGTCGATCTTCTGCCATCTGCGACAGACTAATGTTGAAGTCTTCCCTCAGGTGGAA

ubiquitin

S K I Q D K E G I P P D Q Q R L I F A G K Q L E D G R T L S D Y N I Q K E S T L

BamHI

ACATCTTGTC TAAGGCTAAGAGGTGGTAGGCACGGATCCGGCATCATGGTTCGACCATTGAAGTGCATCGTCGCCGTGCCAAAATATGGGGATTGGCAAGAACGGAGACCTACCCTG 3960  
TGTAGAACAGATTCCGATTTCCACCATCCGTGCC TAGGCCGTAGTACCAAGCTGGTAAC TTGACGTAGCAGCGGCACAGGGTTTTATACCCCTAACCGTTCTTGCTCTGGATGGGAC

ubiquitin

N

DHFR ts degrpm

H L V L R L R G G R H G S G I M V R P L N C I V A V S Q N M G I G K N G D L P W

Scal

GCCTCCGCTCAGGAACGAGTTCAAGTACTTCCAAGAATGACCACAACCTCTTCAGAGGAAGGTAACAGAATCTGGTGATTATGGGTAGGAAAACCTGGTTCTCCATTCTGAGAAGAA 4080  
CGGAGGCGAGTCCCTTGCTCAAGTTCATGAAGGTTCTTACTGGTGTGGAGAAGTCTCCTTCCATTTGCTTAGACCACTAATACCCATCCTTTTGACCAAGAGGTAAGGACTCTTCTT

DHFR ts degrpm

P P L R N E F K Y F Q R M T T T S S E E G K Q N L V I M G R K T W F S I P E K N

TCGACTTTTAAAGGACAGAATTAATATAGTTCTCAGTAGAGAAGCTCAAAGAACCACCACGAGGAGCTCATTCTTTCGCAAAAAGTTTGGATGATGCCTTAAGACTTATTGAACAACCGGA 4200  
AGCTGAAAATTTCTGTCTTAATTATATCAAGAGTCATCTCTTGAGTTCTTGGTGGTGTCTCTCGAGTAAAAGAACGGTTTTCAAACCTACTACGGAATCTGAATAACTTGTGGCCT

DHFR ts degrpm

R L L K D R I N I V L S R E L K E P P R G A H F L A K S L D D A L R L I E Q P E

ATTGGCAAGTAAAGTAGACATGGTTTGGATAGTCGGAGGCAGTCTGTTTTACCAGGAAGCCATGAATCAACCAGGCCACCTCAGACTCTTTGTGACAAGGATCATGCAGGAATTTGAAAG 4320  
TAACCGTTCATTTTCATCTGTACCAAACCTATCAGCCTCCGTCAAGACAAATGGTCTTCGGTACTTAGTTGGTCCGGTGGAGTCTGAGAAACACTGTTCTTAGTACGTCTTAAACTTTC

DHFR ts degrpm

L A S K V D M V W I V G G S S V Y Q E A M N O P G H L R L F V T R I M Q E F E S

TGACACGTTTTTCCCAGAAATTGATTTGGGGAAATATAAACTTCTCCCAGAAATACCCAGGCGTCTCTCTGAGGTCCAGGAGGAAAAAGGCATCAAGTATAAGTTTGAAGTCTACGAGAA 4440  
ACTGTGCAAAAAGGGTCTTTAACTAAACCCCTTATATTTGAAGAGGGTCTTATGGGTCCGAGGAGAGACTCCAGGTCTCTCTTTTCCGTAGTTCATATTCAAACCTTCAGATGCTCTT

DHFR ts degrpm

D T F F P E I D L G K Y K L L P E Y P G V L S E V Q E E K G I K Y K F E V Y E K

GAAAGACGGTACCTACCCATACGATGTTCCAGATTACGCTGCATTTTTAGGAGGGCAAGCCTCGAGGGTAGATCTGGTGCCCGGGGTGGTGTGGTGCCGGAGCCGGTGTGGTGTGAA 4560  
CTTCTGCCATGGATGGGTATGCTACAAGGTCTAATGCGACGTAATAATCTCCCGTTCGGAGCTCCCATCTAGACCACGGGCCACCACGACCACGGCTCGGCCACGACCACGACTT

Acc651  
KpnI

AvaI  
XhoI

BglII

AvaI  
SmaI

HindIII

·DHFR ts degrp →

→ HA epitope

← long oligo 2 (reverse) →

sites for inserting n terminus

K D G T Y P Y D V P D Y A A F L G G Q A S R V D L V P G V V L V P E P V L V L K

GCTTTAATGAGTAAAGGAGAAGAAGCTTTTCACTGGAGTTGTCCCAATCTTGTGTTGAATTAGATGGTGTGTTAATGGGCACAAATTTCTGTCAAGTGGAGAGGGTGAAGGTGATGCAACA 4680  
CGAAATACTCATTCTCTTCTTGAAGAGTACCTCAACAGGGTAAAGAACAACCTAATCTACCACTACAATTACCCGTGTTTAAAGACAGTCACTCTCCCACTTCCACTACGTTGT

→  
·sites

L . . V K E K N F S L E L S Q F L L N . M V M L M G T N F L S V E R V K V M Q H

NcoI

NdeI

TACGGAAAACCTTACCCTTAAATTTTTCACACTACTGAAAACCTACCTGTTCCATGGCCAACACTTGTCACTACTTTCTCTTATGGTGTTCATGCTTTTCAAGATACCCAGATCATATG 4800  
ATGCCTTTTGAATGGGAATTTAAATAAACGTGATGACCTTTTGTGATGGACAAGGTACCGGTTGTGAACAGTGTGAAAGAGAATACCACAAGTTACGAAAAGTTCATGGGTCTAGTATAC

T E N L P L N L F A L L E N Y L F H G Q H L S L L S L M V F N A F Q D T Q I I .

AAGCGGCACGACTTCTTCAAGAGCGCCATGCCTGAGGGATACGTGCAGGAGAGGACCATCTTCTTCAAGGACGACGGGAACACAAGACACGTGCTGAAGTCAAGTTTGAGGGAGACACC 4920  
TTGCGCGTGCTGAAGAAGTTCTCGCGGTACGGACTCCCTATGCACGTCCTCTCCTGGTAGAAGAAGTTCTGCTGCCCTTGATGTTCTGTGCACGACTTCAGTTCAAACCTCCCTCTGTGG  
S G T T S S R A P C L R D T C R R G P S S S R T T G T T R H V L K S S L R E T P

Clal

CTCGTCAACAGGATCGAGCTTAAGGGAATCGATTTC AAGGAGGACGGAAACATCCTCGGCCACAAGTTGGAATACAAC TACAAC TCCCACAACGTATACATCATGGCCGACAAGCAAAAG 5040  
GAGCAGTTGCTAGCTCGAATTCCTTAGCTAAAGTTCCCTCGCCTTTGTAGGAGCCGGTGTCAACCTATGTTGATGTTGAGGGTGTGCATATGTAGTACGGCTGTTTCGTTTTTC  
S S T G S S L R E S I S R R T E T S S A T S W N T T T T P T T Y T S W P T S K R

AACGGCATCAAAGCCAACCTCAAGACCCGCCACAACATCGAAGACGGCGCGTGAACCTCGCTGATCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGACAAC 5160  
TTGCGGTAGTTTCGGTTGAAGTCTGGGCGGTGTGTAGCTTCGCCGCCGACGTTGAGCGACTAGTAATAGTTGTTTATGAGGTTAACCGCTACCGGGACAGGAAAATGGTCTGTTG  
T A S K P T S R P A T T S K T A A C N S L I I I N K I L Q L A M A L S F Y Q T T

PvuII

CATTACCTGTCCACACAATCTGCCCTTTCGAAGATCCCAACGAAAAGAGAGACCACATGGTCCCTCTTGAGTTTGTAAACGCTGCTGGGATTACACATGGCATGGATGAAC TATACAAA 5280  
GTAATGGACAGGTGTGTTAGACGGGAAAGCTTTCTAGGGTTGCTTTTCTCTGGTGTACCAGGAAGAACTCAAACATTGTCGACGACCCCTAATGTGTACCGTACCTACTTGATATGTTT  
I T C P H N L P F R K I P T K R E T T W S F L S L . Q L L G L H M A W M N Y T N

Sall

BamHI

SacI

CATGATGAGCTTGTGACGATCCTGAGCTCCAATTCGCCCTATAGTGAGTCGTATTACAATTCAGTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTGGCGTTACCCAAC TTA 5400  
GTACTACTCGAACAGCTGCC TAGGACTCGAGGTTAAGCGGGATATCACTCAGCATAATGTTAAGTGACCGGCAGCAAAATGTTGCAGCACTGACCCCTTTGGGACCGCAATGGGTTGAAT  
M M S L S T D P E L Q F A L . . V V L Q F T G R R F T T S . L G K P W R Y P T .

PvuII

PvuI

ATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGGACGCGCCCTGTAGCGGGC 5520  
TAGCGGAACGTCGTGTAGGGGAAAGCGGTGACCGCATTATCGCTTCTCCGGCGTGGCTAGCGGGAAGGTTGTCAACGCGTCGGACTTACCGCTTACCCTGCGCGGGACATCGCCGC  
S P C S T S P F R Q L A . . R R G P H R S P F P T V A Q P E W R M G R A L . R R

NgoMIV

CATTAAGCGCGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCGCTCCTTTTCGCTTTCTTCCCTTCTTTCTCGCCACGTTTCGCCGGCTTAGCTA 5640  
GTAATTCGCGCCGCCACACCACCAATGCGCGTGCACCTAGCGGATGTGAACGGTCCGCGGATCGCGGGCAGGAAAGCGAAAGAAGGGAAGGAAAGAGCGGTGCAAGCGCCGAATCGAT

long oligo 1 (forward)

ura4+

I K R G G C G G Y A Q R D R Y T C Q R P S A R S F R F L P F L S R H V R R L S Y

NsiI

CAAATCCCCTGGCTATATGTATGCATTTGTGTTAAAAAGTTTGTATAGATTATTTAATCTACTCAGCATTCTTTCTCTAAATAGGAATTTGTTACTTAATGGAGAAAAAATGTTTCG  
GTTTAGGGTGACCGATATACATACGTAAACACAATTTTTTCAAACATATCTAATAAATTAGATGAGTCGTAAGAAAGAGATTTATCCTTAAACAATGAATTACCTCTTTTTTACAAGC

5760

ura4+

K S H W L Y V C I C V K K V C I D Y L I Y S A F F L . I G I C Y L M E K K M F R

ATTTACCTAGTGTATTTGTTTGTATACTCAGTTTAAATTTCAAACATCCATTCTATCTTGTGTAATTTTGGCATGGTGA AAAAGATAATCAGCCTTATAATCTTTACAAAAGTAAGAAA  
TAAATGGATCACATAAACAAACATATGAGTGCAAATTAAGTTTGTAGGTAAGATAGAACACATTA AAAACCGTACCCTTTTTCTATTAGTCGGAATATTAGAAATGTTTTATTCTTT

5880

ura4+

F T . C I C L Y T H V . F Q T S I L S C V I F G M V K K I I S L I I F T K V R N

TTCTGTAATAAGCCTTAATGCCCTTGCTTTAAATTA AAAAGTTCTTTTTCATGATAATGTTTGCACCTTGTGAATATTTTTAGATAGTTCTGTGAGGTATAATTAAGATGTTTTAGA  
AAGACATTTATTCGGAATTACGGGAACGAAATTTAATTTTACCAAGAAAAAGTACTATTACAAACGTGAAACACTTATATAAAATCTATCAAGACTCCATATTAATTTACAAAATCT

6000

ura4+

S V N K P . C P C F K L K W F F F M I M F A L C E Y I L D S S V R Y N . D V L E

GACTTATACAATTTTGTCTTTATAAATCTTAATTTGATTTTACCATCCCAGTTTAACTATGCTTCGTCGGCATCTCTGCACATGTCGTGTTTTCTTACCGTATTGCCTACCAAGAACCT  
CTGAATATGTTAAAACAGAAATATTTAAGAATTAAC TAAAATGGTAGGGTCAAATTTGATACGAAGCAGCCGTAGAGACGTGTACAGCACAAAAGAATGGCATAACAGGATGGTTCTTGGA

6120

ura4+

T Y T I L S L . I L N . F Y H P S L T M L R R H L C T C R V F L P Y C P T K N L

CTTTTTGCTTGGATCGAAATTAAGGTTTAAAAGCAAAGTTATGGATGCTAGAGTATTTCAAAGCTATTCAGCTAGAGCTGAGGGGATGAAAAATCCCATTGCCAAGGAATTGTTGGCT  
GAAAAACGAACCTAGCTTTAATTTCCAATTTTCGTTTCAATACCTACGATCTCATAAAGTTTCGATAAGTCGATCTCGACTCCCTACTTTTTAGGGTAACGGTTCCCTAACACCCGA

6240

ura4+

F F A W I E I K G L K A K L W M L E Y F K A I O L E L R G . K I P L P R N C W L

TTGATGGAAGAAAAGCAAAGCAACTTGTCACTCGCGGTCGATTTGACGAAGAAATCCGAAATCTTAGAATGGTAGATAAAATTTGGACCTATGCTGTGTTATCAAGACACATATTGAC  
AACTACCTCTTTTCGTTTCGTTGAACAGTCAGCGCCAGCTAACTGCTTCTTTAGGCTTTAGAATCTTAACCATCTATTTAACCTGGGATACAGACACAATAGTTCTGTGTATAACTG

6360

ura4+

. W K K S K A T C Q S R S I . R R N P K S . N W . I K L D P M S V L S R H I L T



GTTGTCGAGGATTTTCGACCAGGATATGGTAGAAAACTGGTGGCCTTAGGTAAAAAGCATCGTTTTCTTATCTTTGAGGATCGCAAATTCGCAGACATTGGAAATACCGTCAAGCTACAA  
CAACAGCTCCTAAAGCTGGTCCATACCATCTTTTGGACCACCGGAATCCATTTTCGTAGCAAAAAGAAATAGAAACTCCTAGCGTTTAAAGCGTCTGTAACCTTTATGGCAGTTCGATGTT

6480

ura4+

L S R I S T R I W . K N W W P . V K S I V F L S L R I A N S Q T L E I P S S Y N

Nsil

TATGCATCTGGTGTGTACAAAATGCTTCTGGGCTCATATCACAAATGGCATACAGTGCCAGGCGAGGGTATTATACAAGGCCTCAAAGAAGTTGGTTTACCTTTGGGACGTGGTCTC  
ATACGTAGACCACACATGTTTTAACGAAGAACCAGATAGTGTAAACGGTATGTCACGGTCCGCTCCATAATATGTTCCGGAGTTTCTTCAACCAATGGAAACCTGCACCAGAG

6600

ura4+

M H L V C T K L L L G L I S Q I A I O C O A R V L Y K A S K K L V Y L W D V V S

TTGCTTTGGCTGAAATGCTTCCAAAGGCTCTTTGGCTACTGGTTCCTACACAGAGAAAACCTTAGAATGGTTTGAAGAAGCATACCGATTTTGGCTTTATAGCTGGTCGTCGA  
AACGAAAACCGACTTTACAGAAGGTTCCGAGAAACCGATGACCAAGGATGTGTCTTTTGAATCTTACCAAACCTCTCGTATGGCTAAAAACGAAACGAAATATCGACCAGCAGCT

6720

ura4+

C F W L K C L P K A L W L L V P T Q R K P . N G L R S I P I F A L A L . L V V D

TTTCTAACCTTCAAAGCGACTACATAACTATGCCCTGGTATCGGCTTGGATGTTAAAGGAGACGGGCTGGGACAGCAATATCGTACTCCTGAAGAAGTGATTGTAACTGCGGTAGC  
AAAGGATTGGAAGTTTCGCTGATGATTGATACAGGGGACCATAGCCGAACCTACAATTTCTCTGCCCGACCCTGTCGTTATAGCATGAGGACTTCTTACTAACATTTGACGCCATCG

6840

ura4+

F L T F K A T T . L C P L V S A W M L K E T G W D S N I V L L K K . L . T A V A

EcoRV

GATATCATCATTGTTGGTCGTGGAGTCTATGGAGCTGGTTCGTAATCCTGTTGTGCGAAGCCAGAGATATAGAGAAGCTGGTTGGAAGGCATATCAGCAAAGACTTTCTCAGCATTAAAA  
CTATAGTAGTAACAACCGACCTCAGATACCTCGACCAGCATTAGGACAACAGCTTCGGTCTCTATATCTTTCGACCAACCTCCGTATAGTCGTTTCTGAAAGAGTCGTAATTTT

6960

ura4+

I S S L L V V E S M E L V V I L L S K P R D I E K L V G R H I S K D F L S I K K

Scal

AAGACTAATGTAATAATTTTTGGTTGGTATTGAAAAAGTCGATGCCTTGTTCGCTTTGTTTTCTAGGCGTTTTATGTCAGAAGGCATTTAGAATTAGTATACAAGTACTCTTTGGT  
TTCTGATTACATTTTAAAAAACCAACCAATAACTTTTTCAGCTACGGAACAAACGCAAAACAAAGGATCCGCAAAATACAGTCTCCGTAATCTTAATCATATGTTTCATGAGAAACCA

7080

ura4+

R L M . N F F G W L L K K S M P C L R L F S . A F Y V R R H L E L V Y K Y S L V

SspI

AAAATTTTATGTAGCGACTAAAATATTAAC TATTATAGATAAACACCTTGGGAATAAAAAGTAA TTTGCTATAGTAATTTATTAACATGCTCCTACAACATTACCACAATCTTTTCTCT  
7200  
TTTTAAAATACATCGCTGATTTTATAATTGATAATATCTATTTGTGGAACCCCTATTTTTCATTAACGATATCATTAAATAATTTGTACGAGGATGTTGTAATGGTGTAGAAAAGAGA

ura4+

K F Y V A T K I L T I I D K H L G N K K . F A I V I Y . T C S Y N I T T I F S L

TGGATTGACATTGAATAAGAAAAGAGTGAATTTTTTTAGACTTGAATGATAACTATGTACAAAGCCAATGAAAGATGTATGTAGATGAATGTAATAACCATGTAGACAAAACAAGATAA  
7320  
ACCTAACTGTAACCTATTCTTTCTCACTTAAAAAATCTGAACATTACTATTGATACATGTTTCGGTTACTTTCTACATACATCTACTTACATTTTATGGTACATCTGTTTGTCTATT

ura4+

G L T L N K K R V N F F R L V M I T M Y K A N E R C M . M N V K Y H V D K Q D K

NgoMIV

AACTTGGTTATAAACATTGGTGTGGAACAGAATAAATTAGATGTCAAAAAGTTTCGTCAATATCACAAAGCCGGCTTTCCTCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGA  
7440  
TTGAACCAATATTTGTAACCAACCTTGCTTATTATCTACAGTTTTCAAAAGCAGTTATAGTGTTCGGCCGAAAGGGGCAGTTCGAGATTTAGCCCCGAGGGAAATCCCAAGGCT

ura4+

T W L . T L V L E Q N K L D V K K F R Q Y H K P A F P V K L . I G G S L . G S D

TTTAGTGCTTTACGGCACCTCGACCCAAAAAATTTGATTAGGGTGATGGTTCACGTAGTGGCCATCGCCCTGATAGACGGTTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAAT  
7560  
AAATCACGAAATGCCGTGGAGCTGGGGTTTTTGAAC TAATCCCACTACCAAGTGCATCACCCGGTAGCGGACTATCTGCCAAAAAGCGGAAACGCAACCTCAGGTGCAAGAAATTA

L V L Y G T S T P K N L I R V M V H V V G H R P D R R F F A L . R W S P R S L I

AGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTATAAGGGATTTGCCGATTTCCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAA  
7680  
TCACCTGAGAACAAGGTTTGACCTTGTGTGAGTTGGGATAGAGCCAGATAAGAAAAC TAATATTCCTAAAACGGCTAAAGCCGGATAACCAATTTTTTACTCGACTAAATTGTTTTT

V D S C S K L E Q H S T L S R S I L L I Y K G F C R F R P I G . K M S . F N K N

SspI

TTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTTAGGTG  
7725  
AAATTGCGCTTAAAATTGTTTTATAATTGCGAATGTTAAATCCAC

L T R I L T K Y . R L Q F R W